

Detecting Blue Whale Calls By Extending a Speech Processing Algorithm: A Proof of Concept

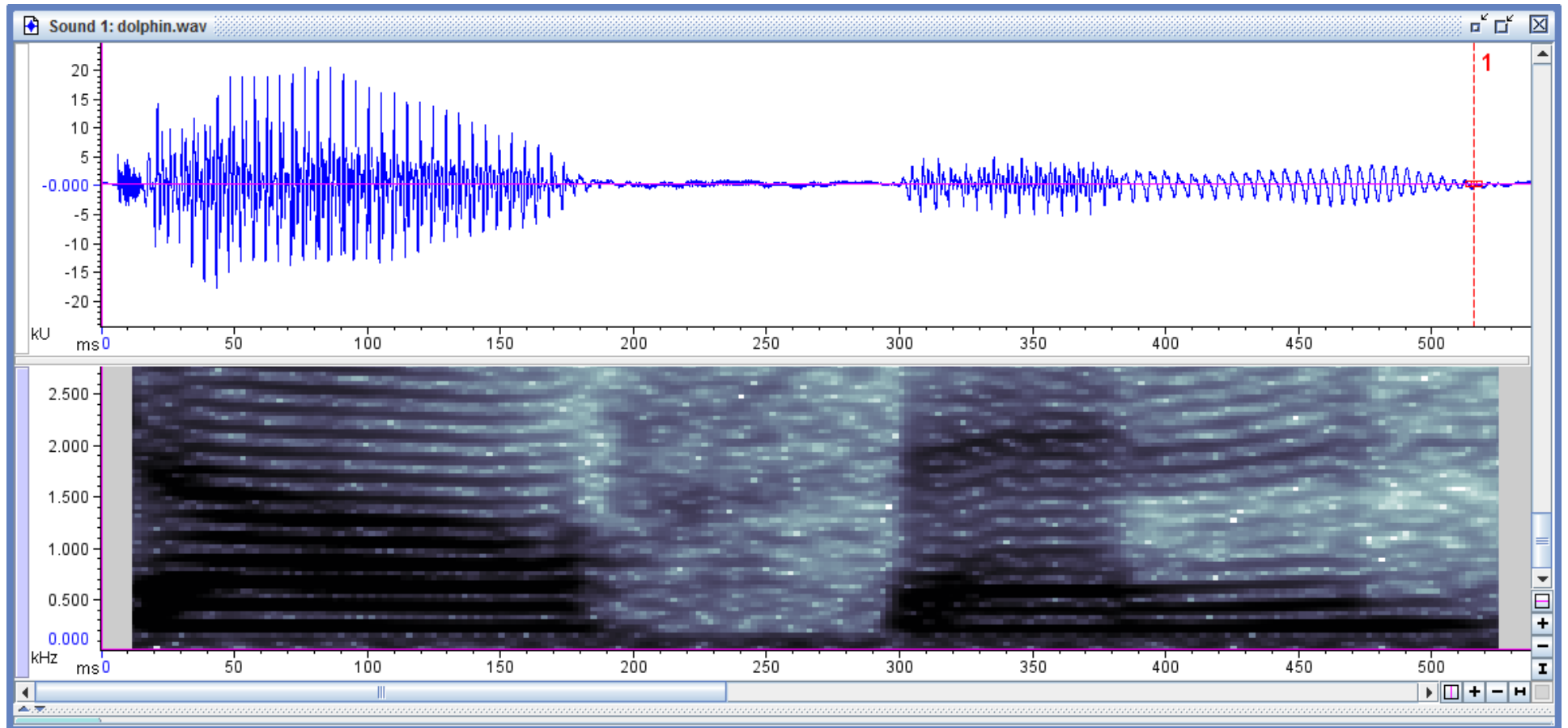
Steve Lockhart
z0 Consulting

DCLDE 2015

Outline

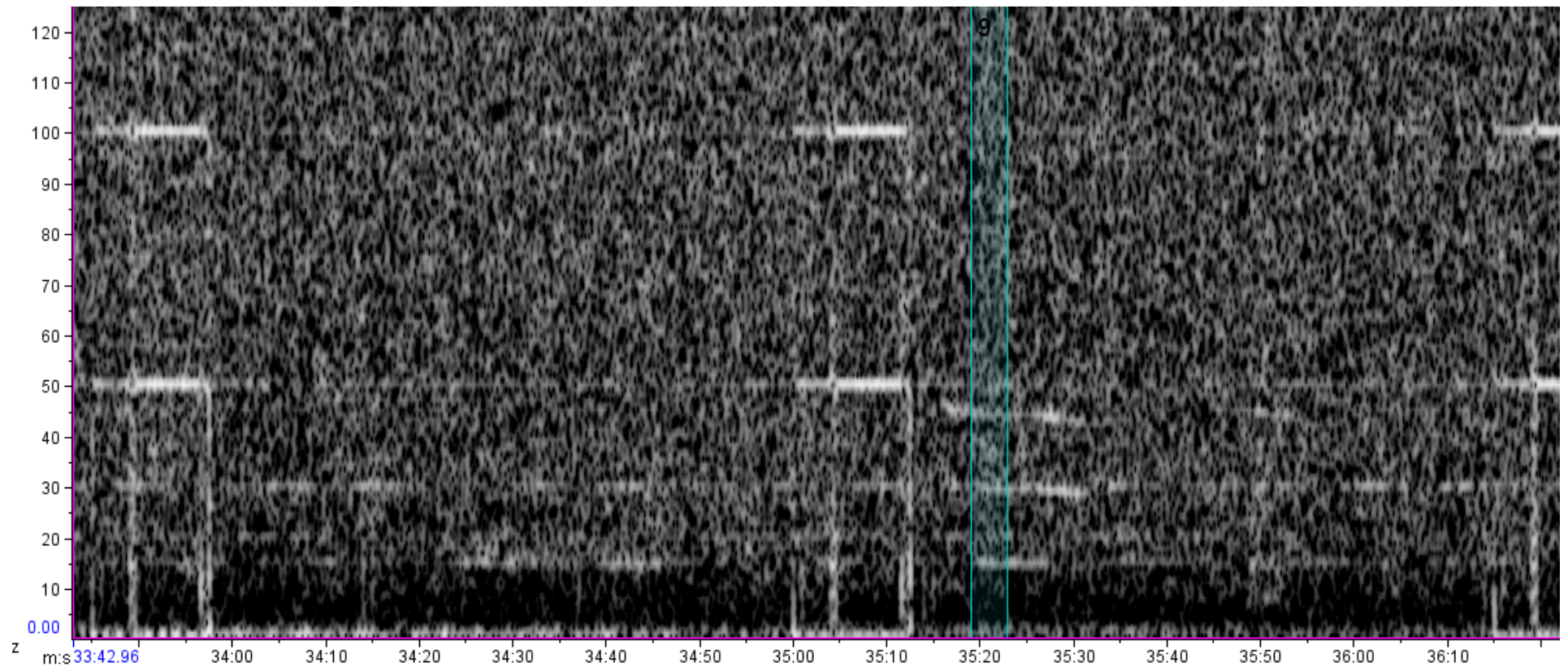
- Approach, Using a Simplified “Synthetic” Signal
- Results, using DCLDE 2015 Training Data Set
 - Automatic Detection of Blue Whale B Calls
 - Temporal Analysis of Detections
- Results, using Another Data Set
 - Automatic Detection of Blue Whale B Calls
 - Compare to Manual Detection
 - Temporal Analysis of Detections
- Potential Benefits of this Approach

Approach: Bioacoustic Signals with Harmonics



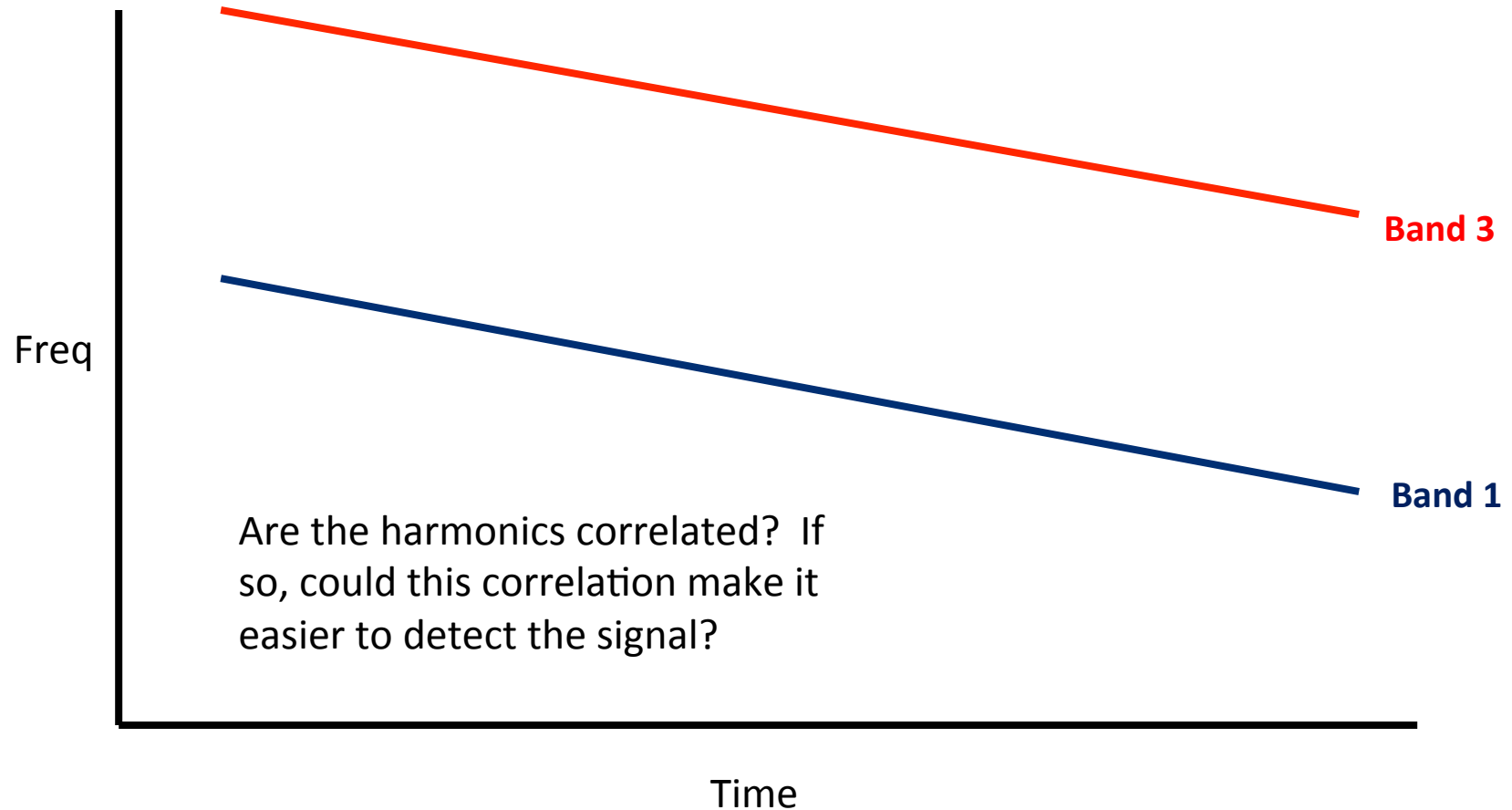
See the harmonic structure of vowels

Approach: Bioacoustic Signals with Harmonics



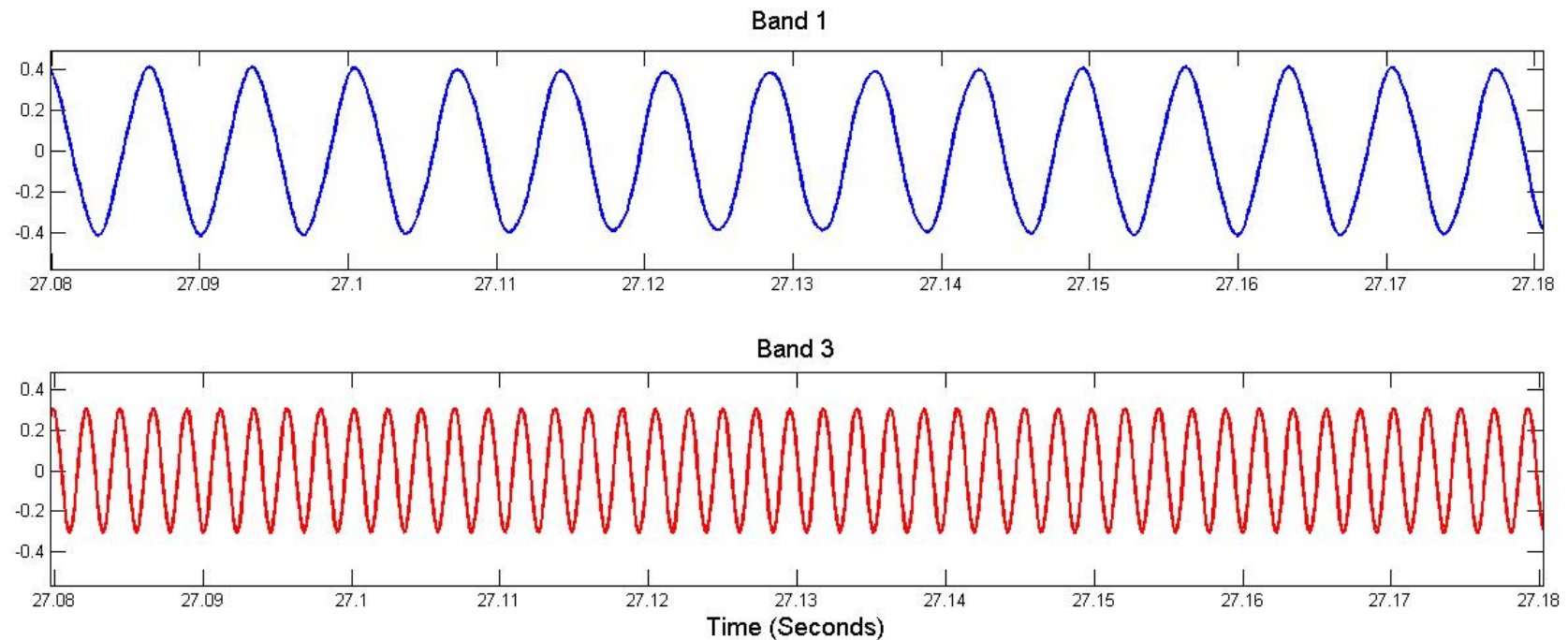
Automatically detected B-Call,
highlighted around 35:20

Approach: Spectrogram of Synthetic Signal

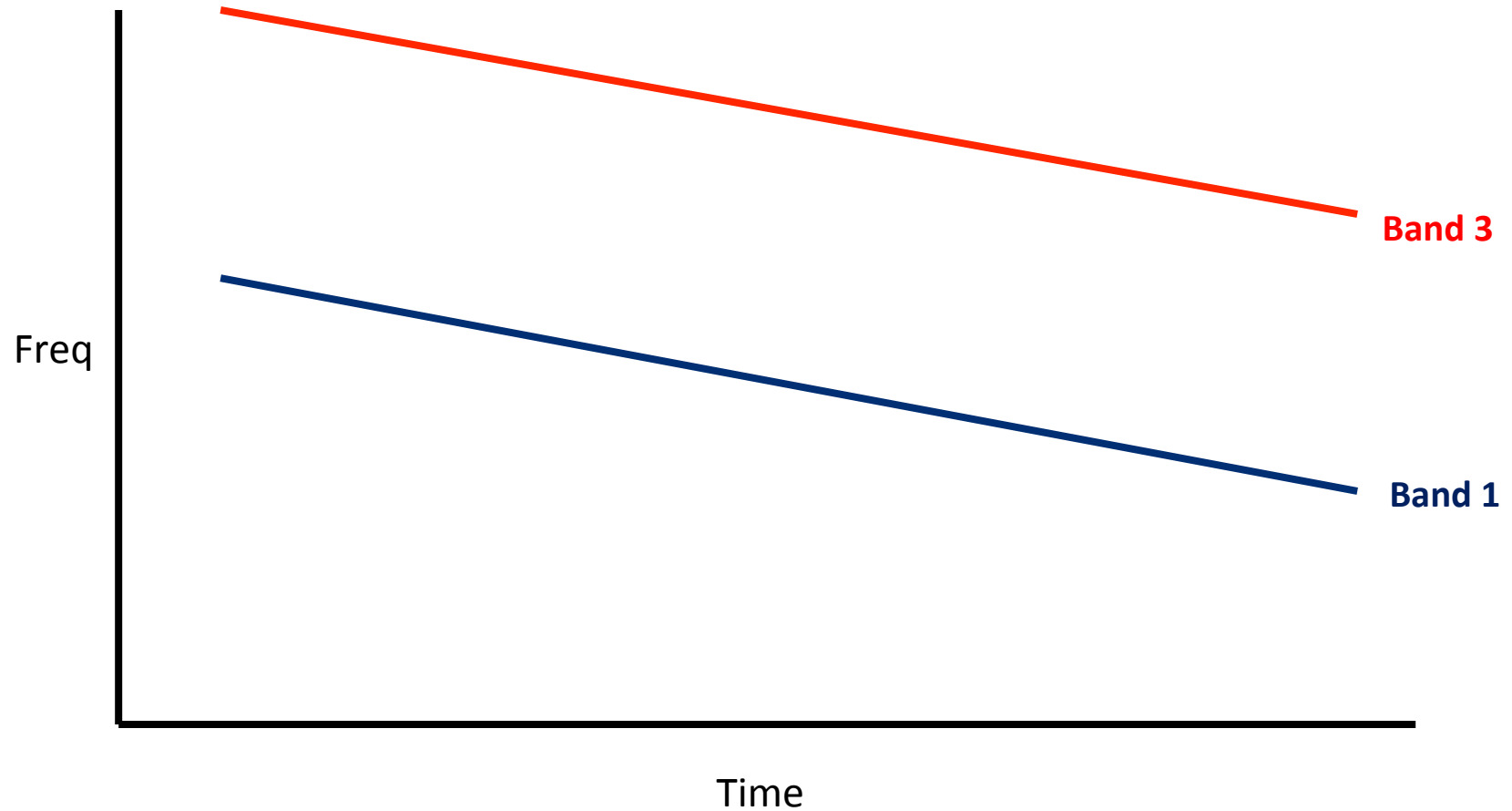


Approach: Bandpass-Filtered Synthetic Signal

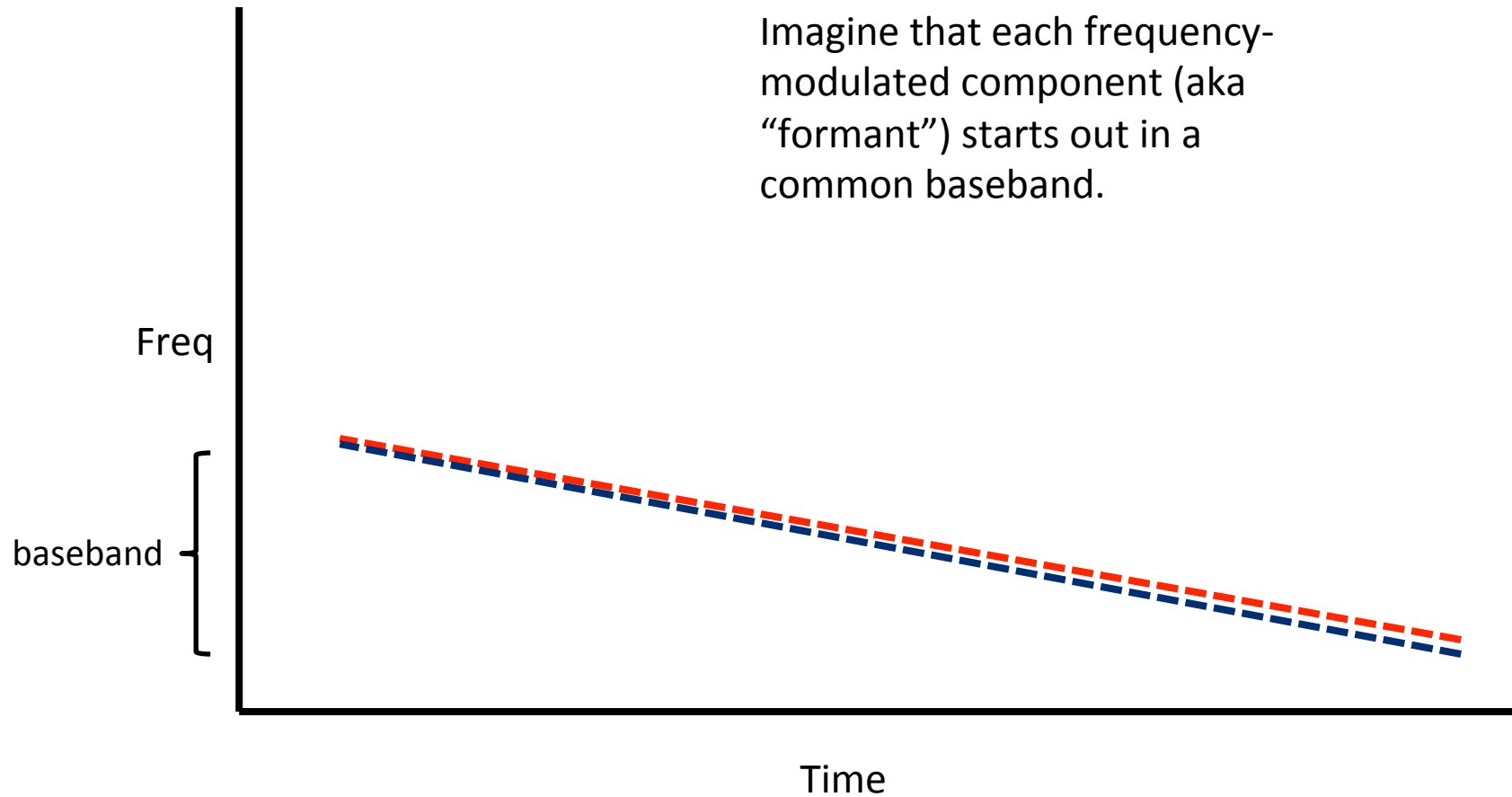
Chirp With Harmonic after Band Pass Filter



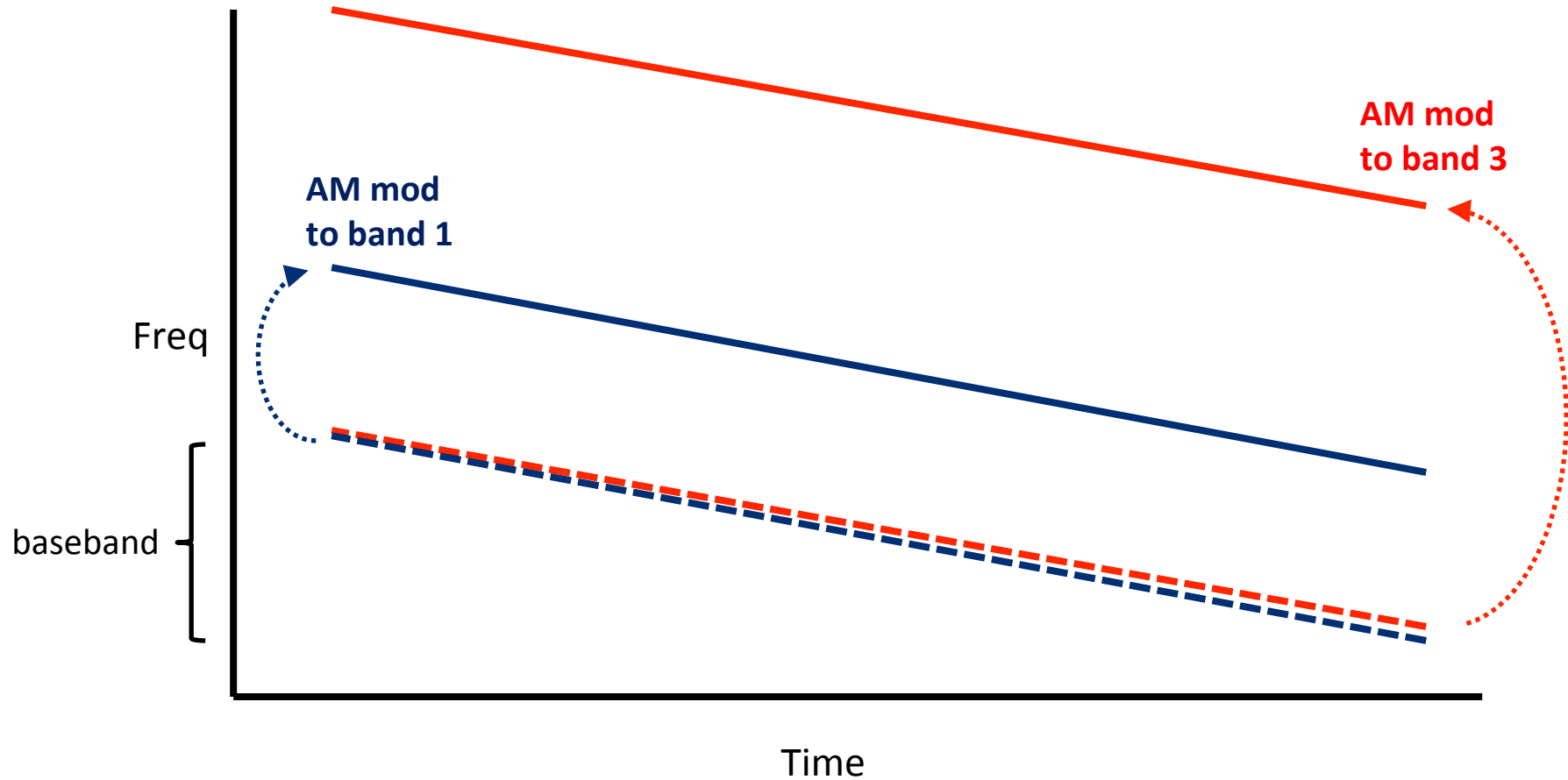
Approach: Spectrogram of Synthetic Signal



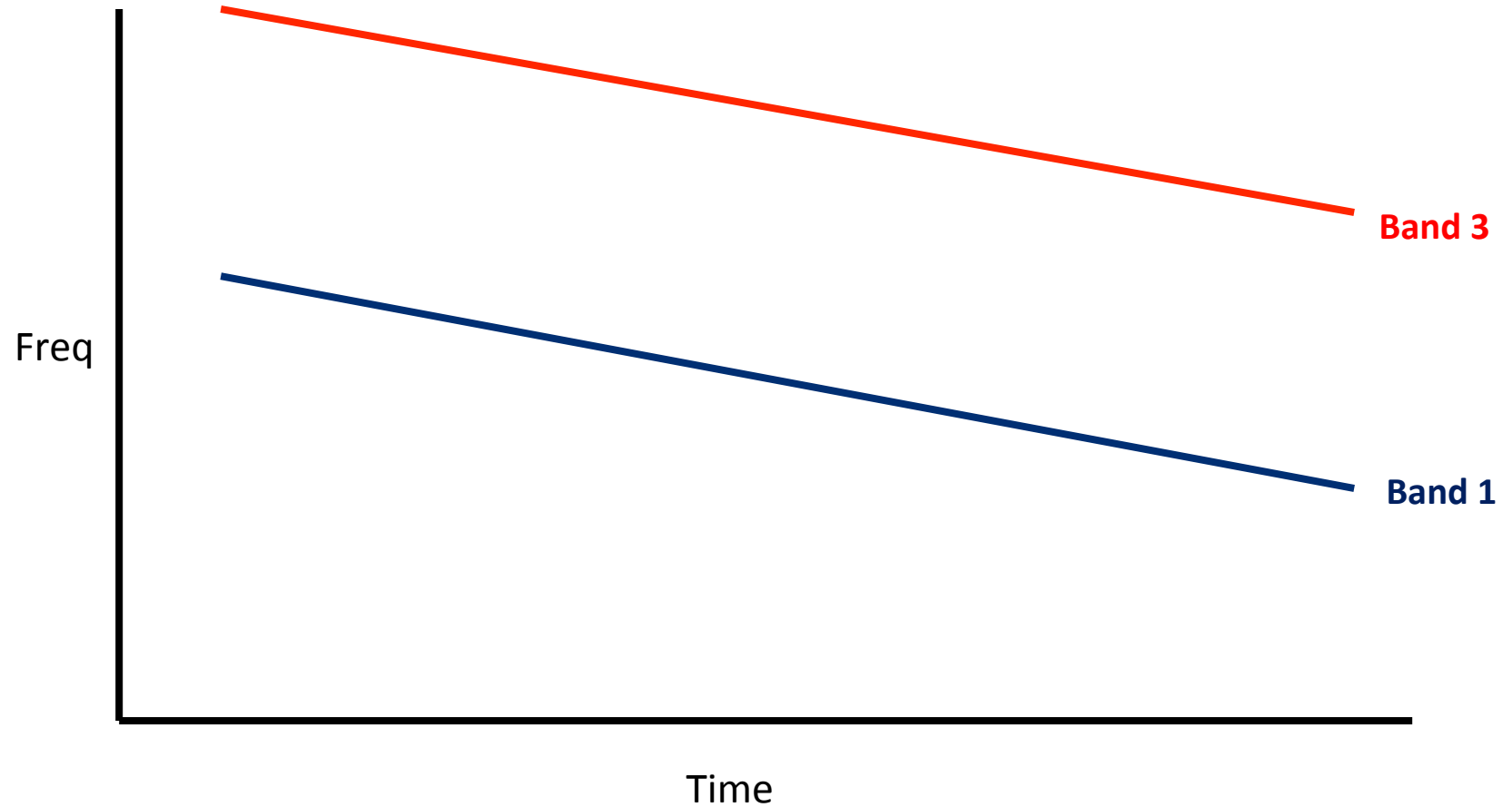
Approach: Model of Formants in a Vowel



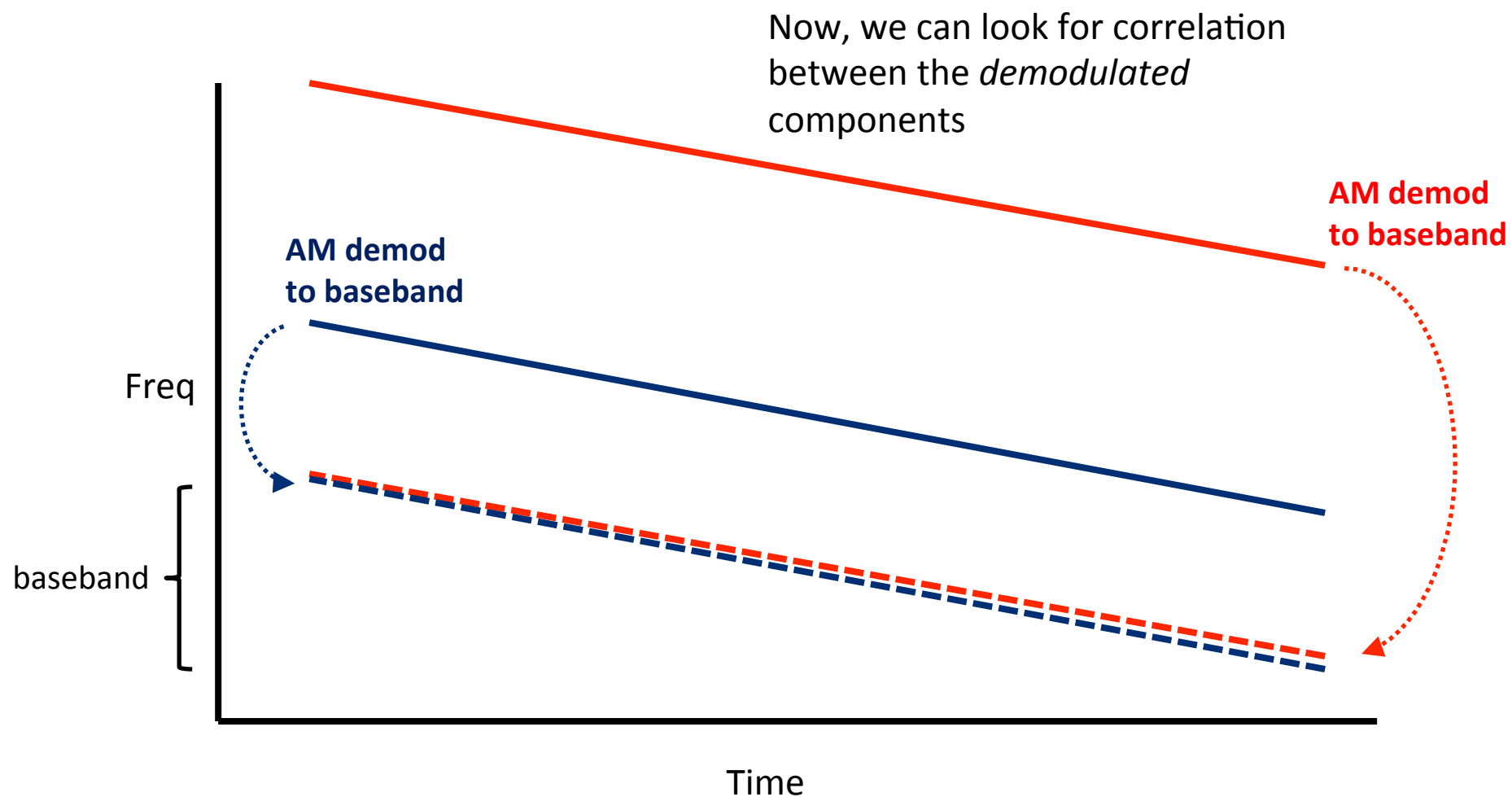
Approach: Model of Formants in a Vowel



Approach: Spectrogram of Synthetic Signal

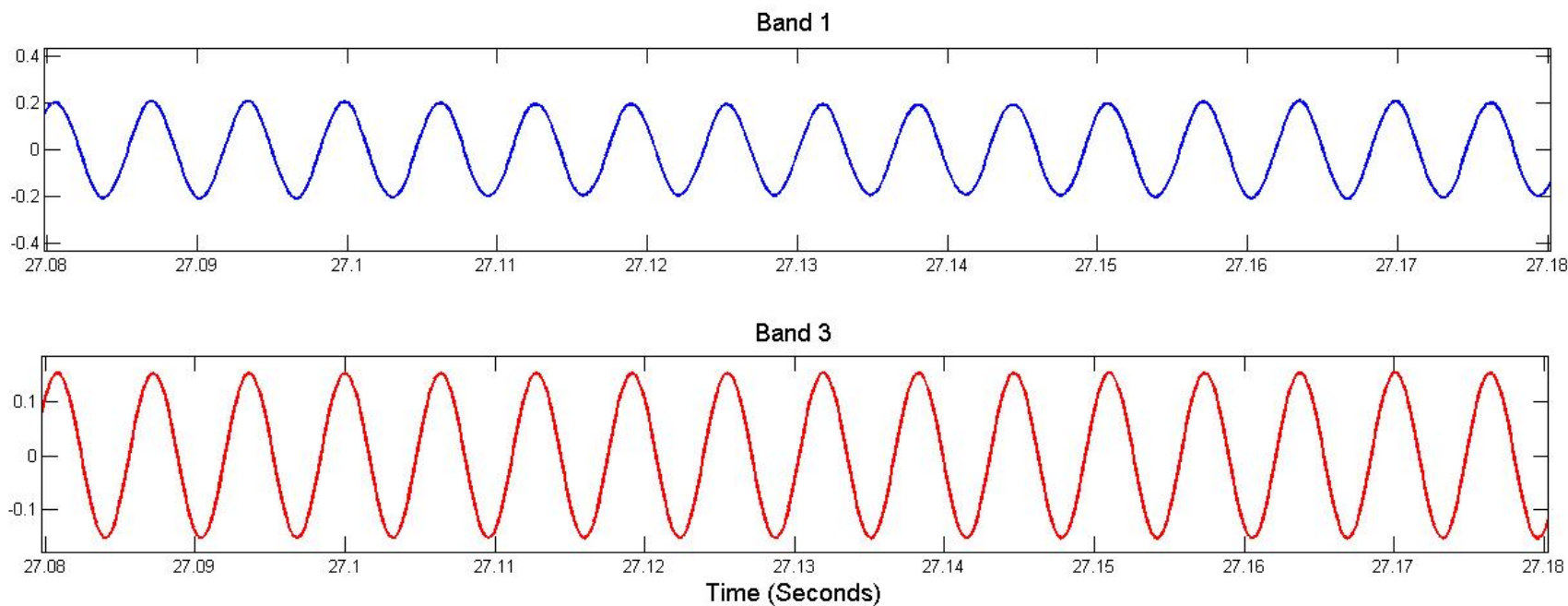


Approach: AM Demod to Get Baseband Components



Approach: BP-Filter and AM Demod

Chirp With Harmonic after Band Pass Filter + AM Demodulation

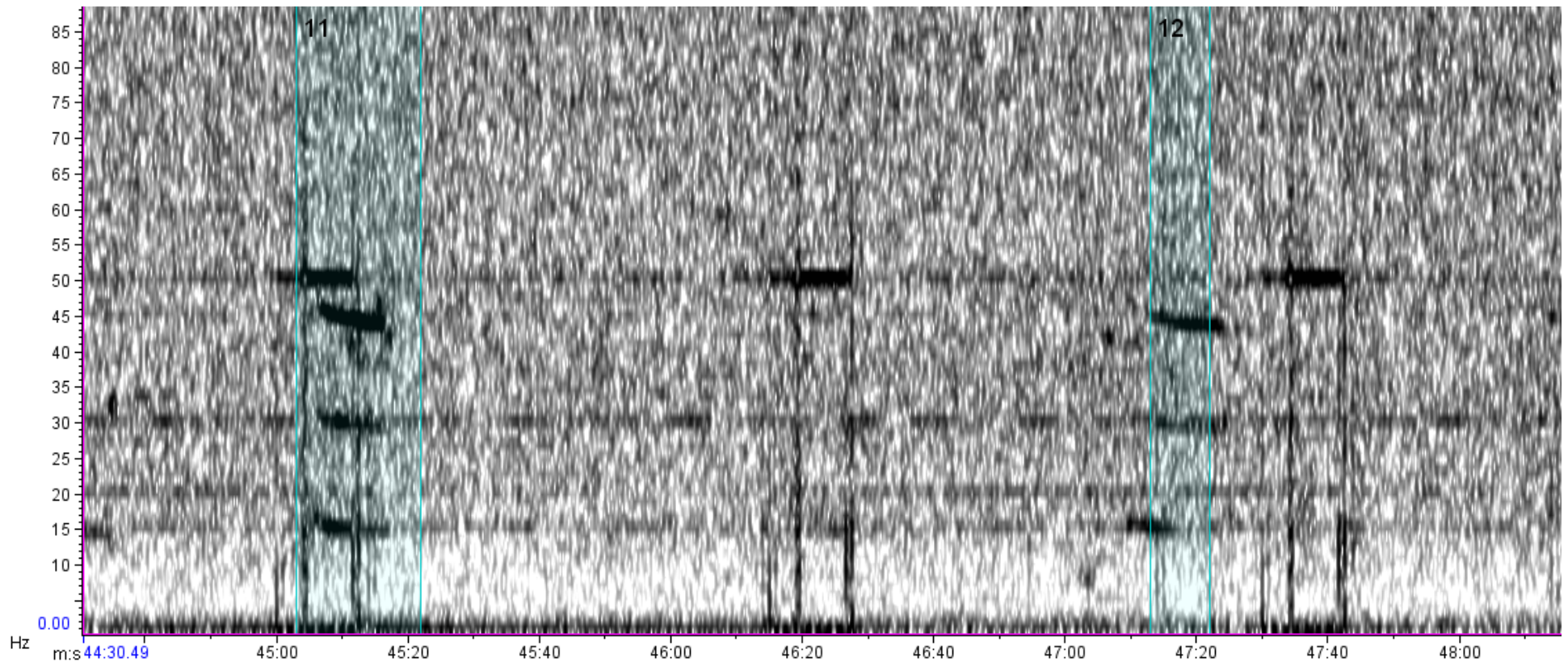


After bandpass filtering and AM demodulation, the demodulated components are well correlated. (Here, I'm zooming in on a subset in time.)

Approach: Summary of Steps

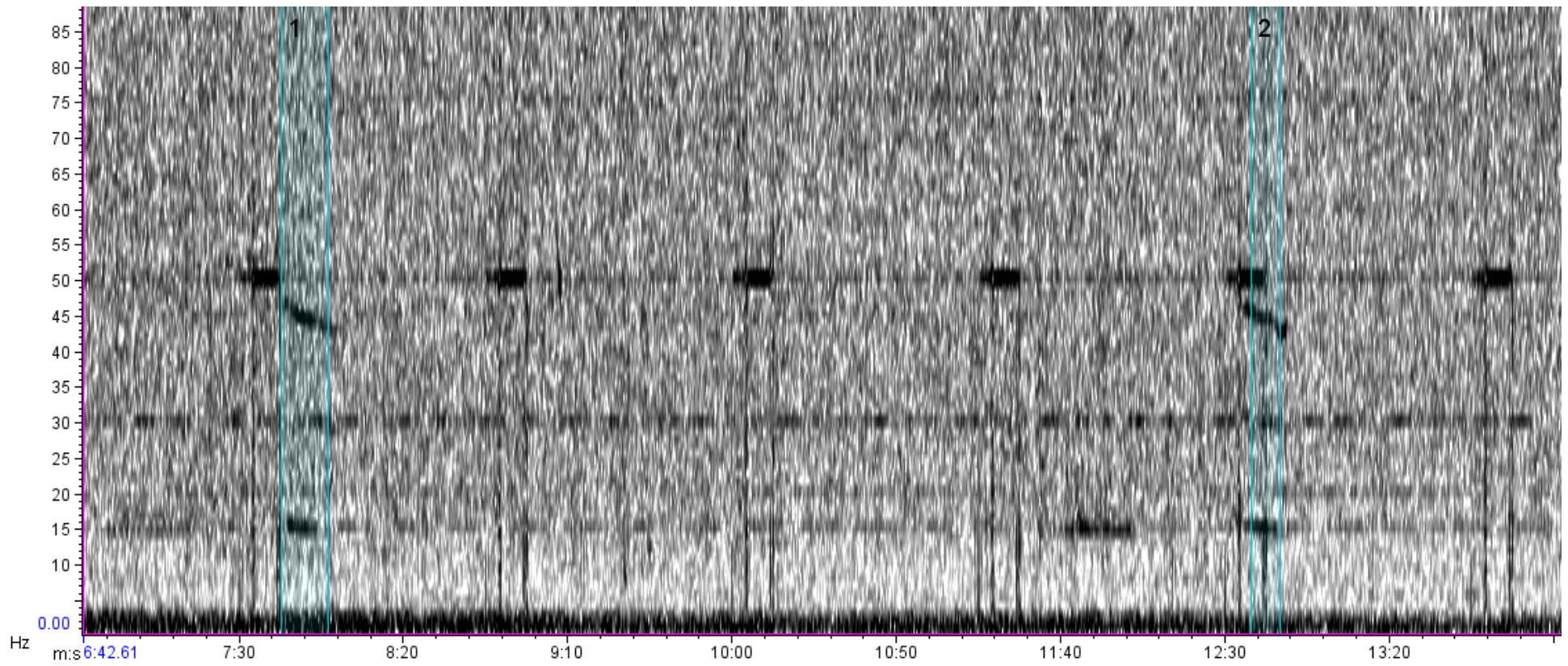
- Define N bands
- Bandpass filter each band
- AM demodulate each component to the baseband.
- For each 50%-overlapping 10-second window in time:
 - Determine the covariance between demodulated component M and demodulated component N.
 - Look for patterns e.g.
 - If the covariance between components 3 and 1 (or 3 and 2) ranks high and is above a threshold, this window contains the signal.
- Fuse together adjacent signal windows to form the “detection events”.
- Import detection events into Raven “selections” for display

Results: Fused Detection Events (B Calls)

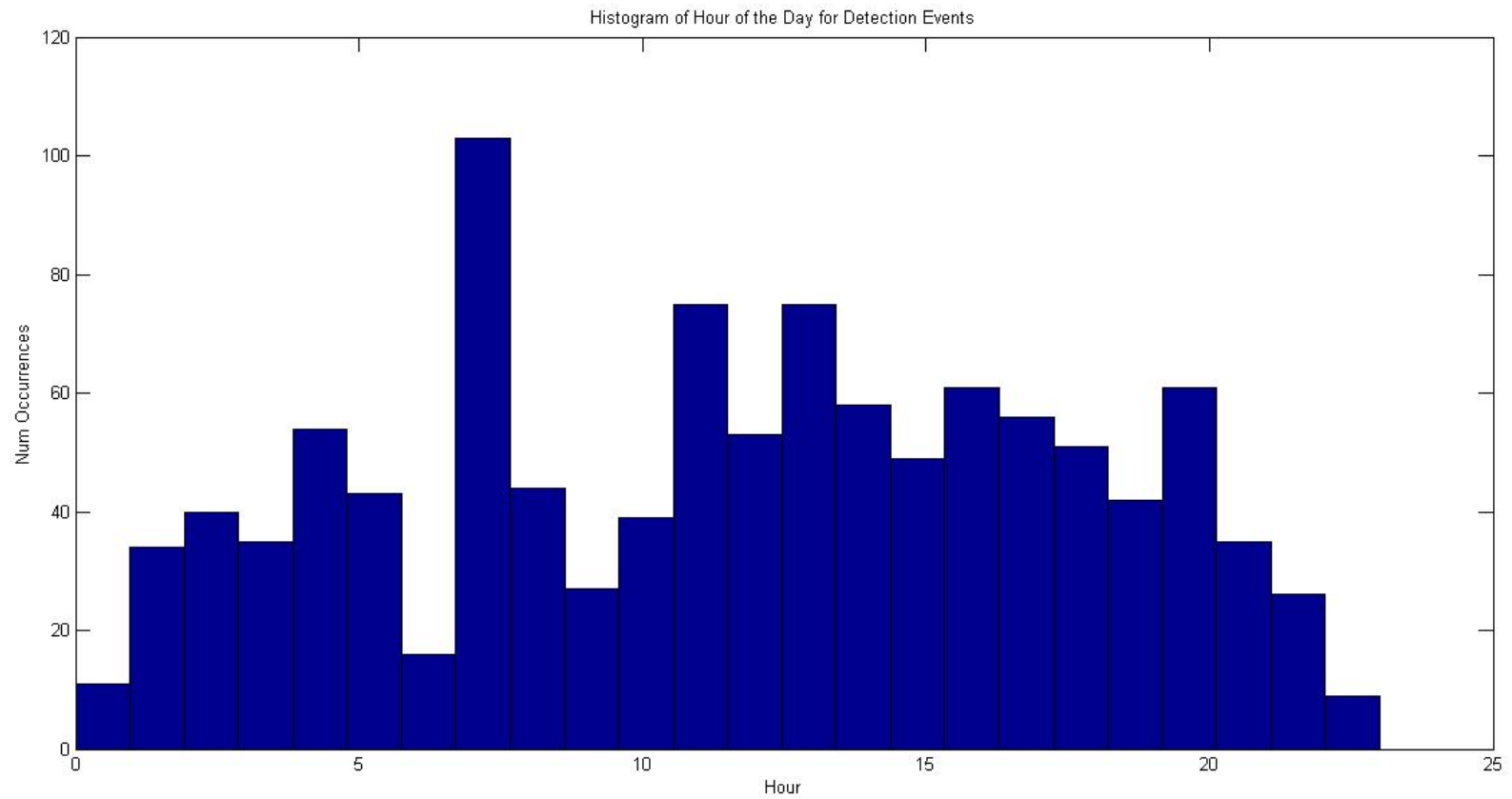


The fused detection events were imported into Raven for display.

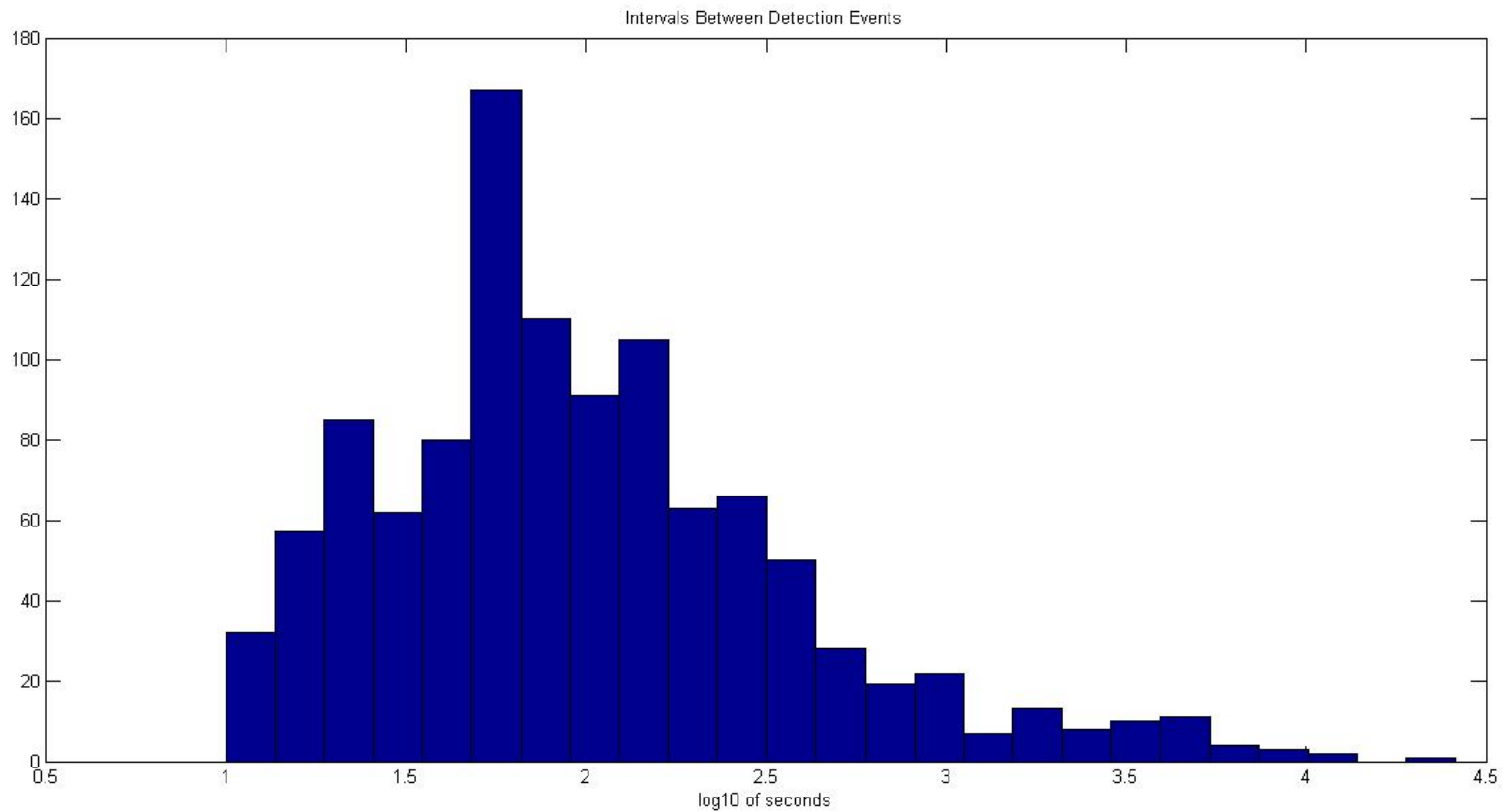
Results: More Fused Detection Events (B Calls)



Results: Time-of-Day for Detected B Calls (DCLDE)

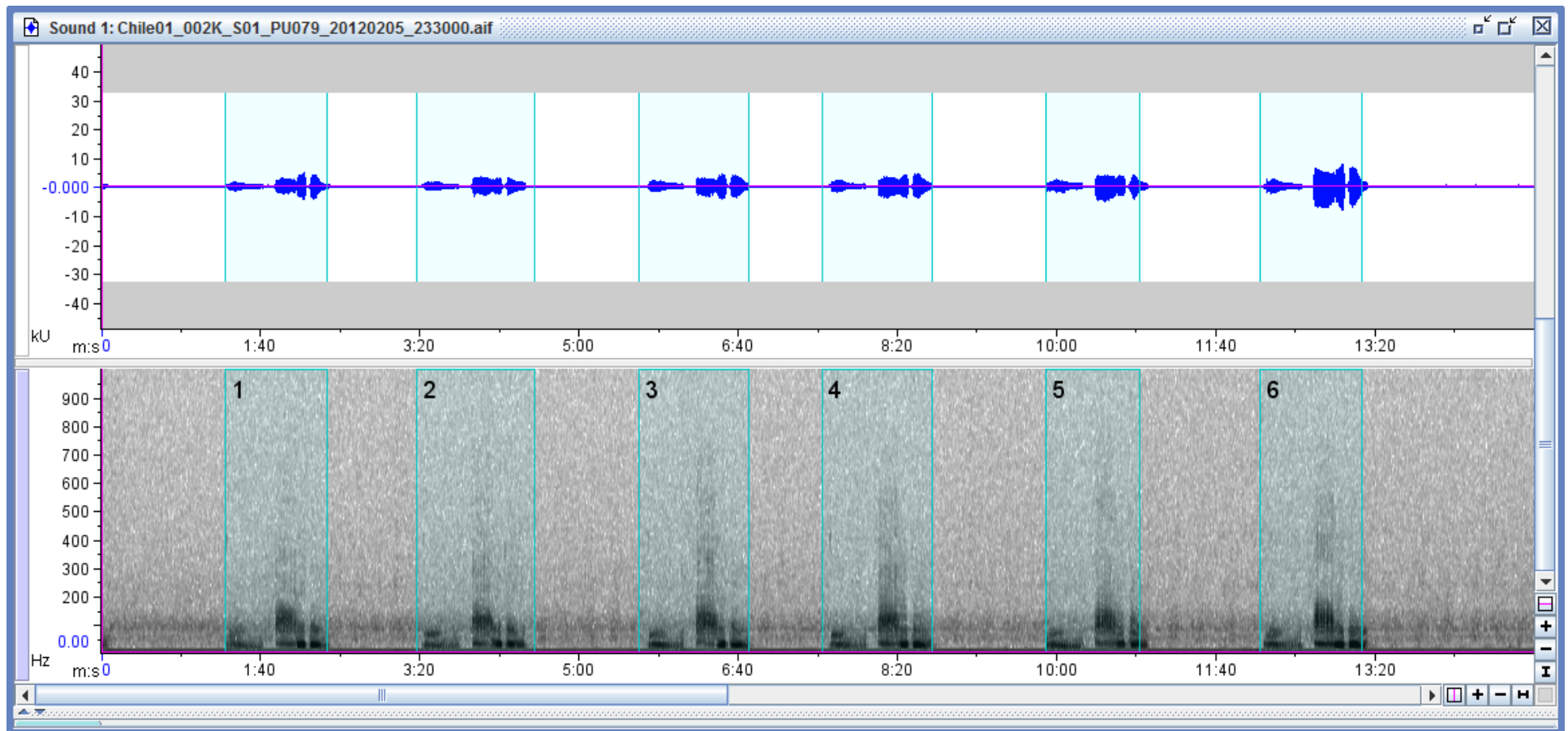


Results: Δt Between Detected B Calls (DCLDE)



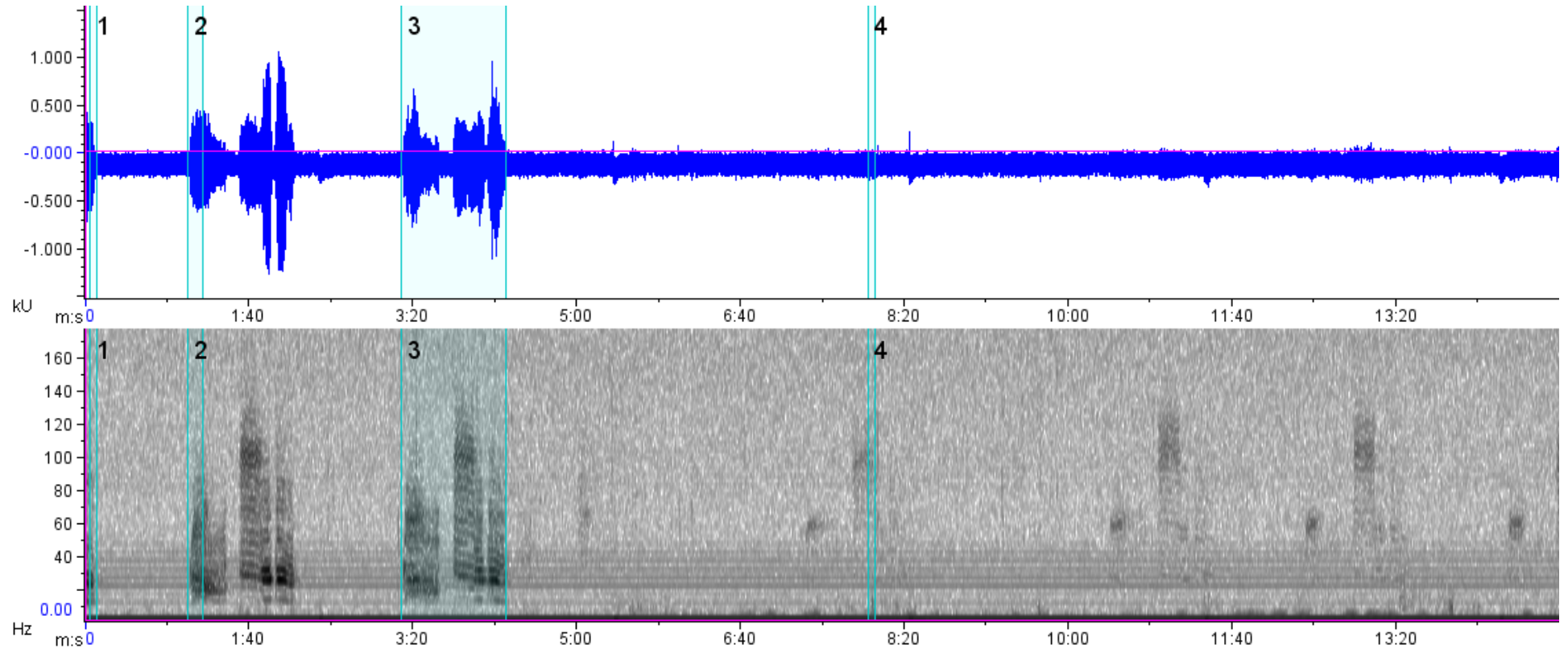
For the DCLDE dataset, the peak of the distribution is around 56 seconds between B Calls.

Results: Fused Detection Events (Chile)



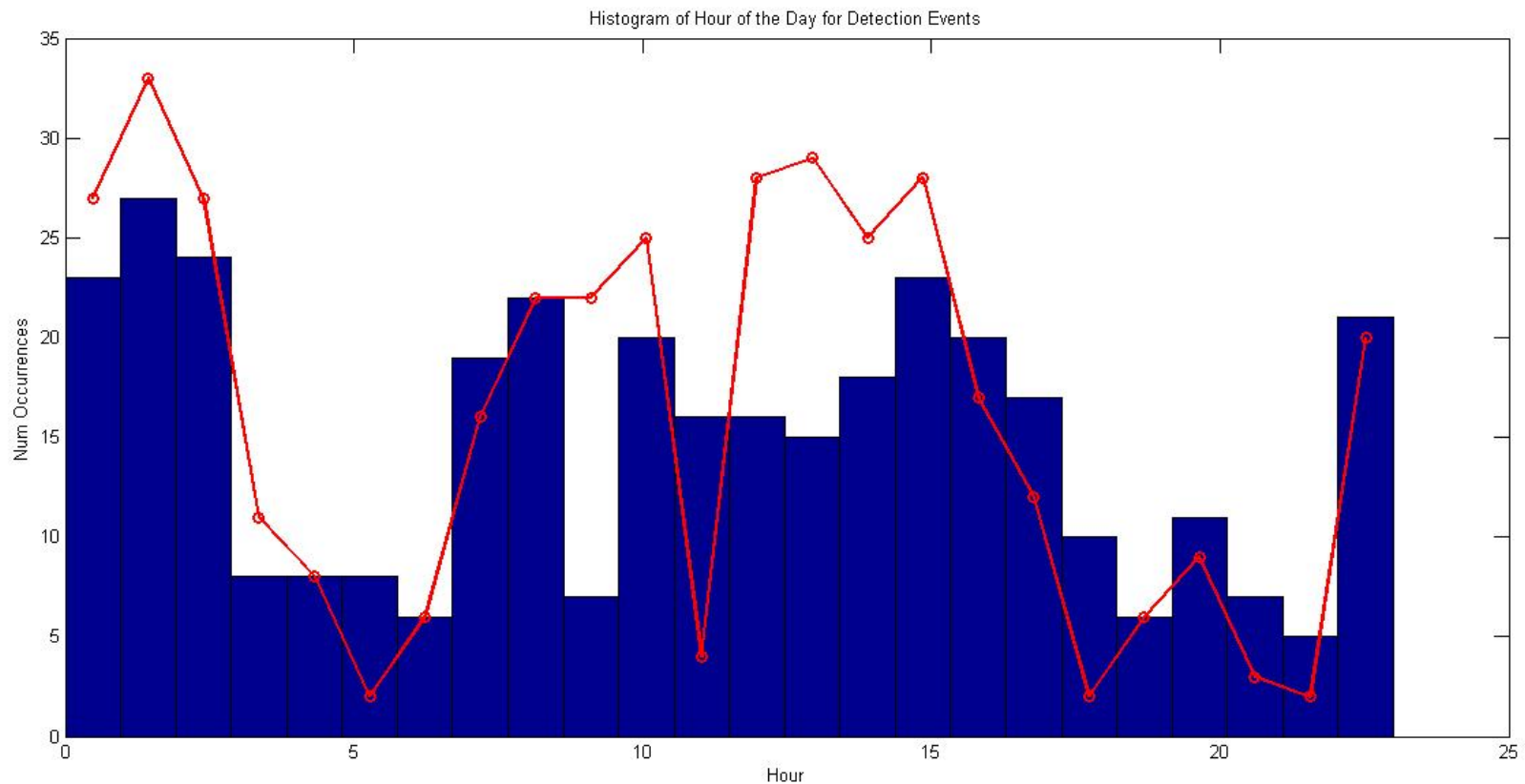
(wav files from Melimoyu Ecosystem Research Institute)

Results: Fused Detection Events (Chile)



(wav files from Melimoyu Ecosystem Research Institute)

Results: Time-of-Day for Detected B Calls (Chile)

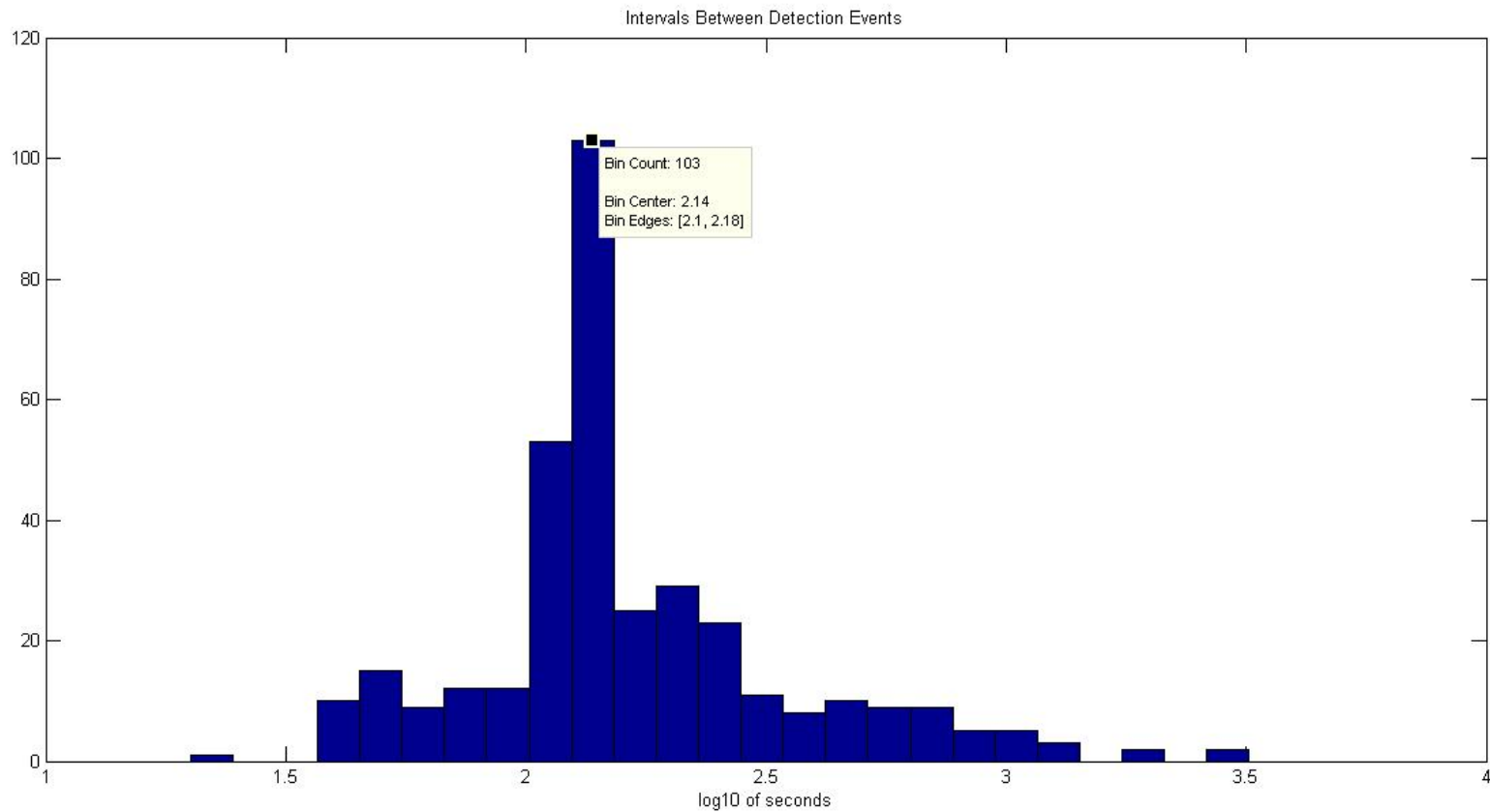


False Alarm Rate = 12%

Miss Rate = 19%

(wav files from Melimoyu Ecosystem Research Institute)

Results: Δt Between Detected B Calls (Chile)



For the dataset from Chile, the peak of the distribution is around 138 seconds between detected B Calls. (wav files from Melimoyu Ecosystem Research Institute)

Compare to Other Approaches

- **Image Processing of Spectrogram** discards phase information, which may be useful for detection.
 - This new approach does not discard phase information.
- Performance of a **Matched Filter** can be degraded by multi-path.
 - This new approach just looks at the covariance between harmonically-related components (after AM demodulation). If the correlation exists along one path, it may also be there along another path.

Acknowledgements

I'd like to thank the team at Scripps for putting together the DCLDE data sets.

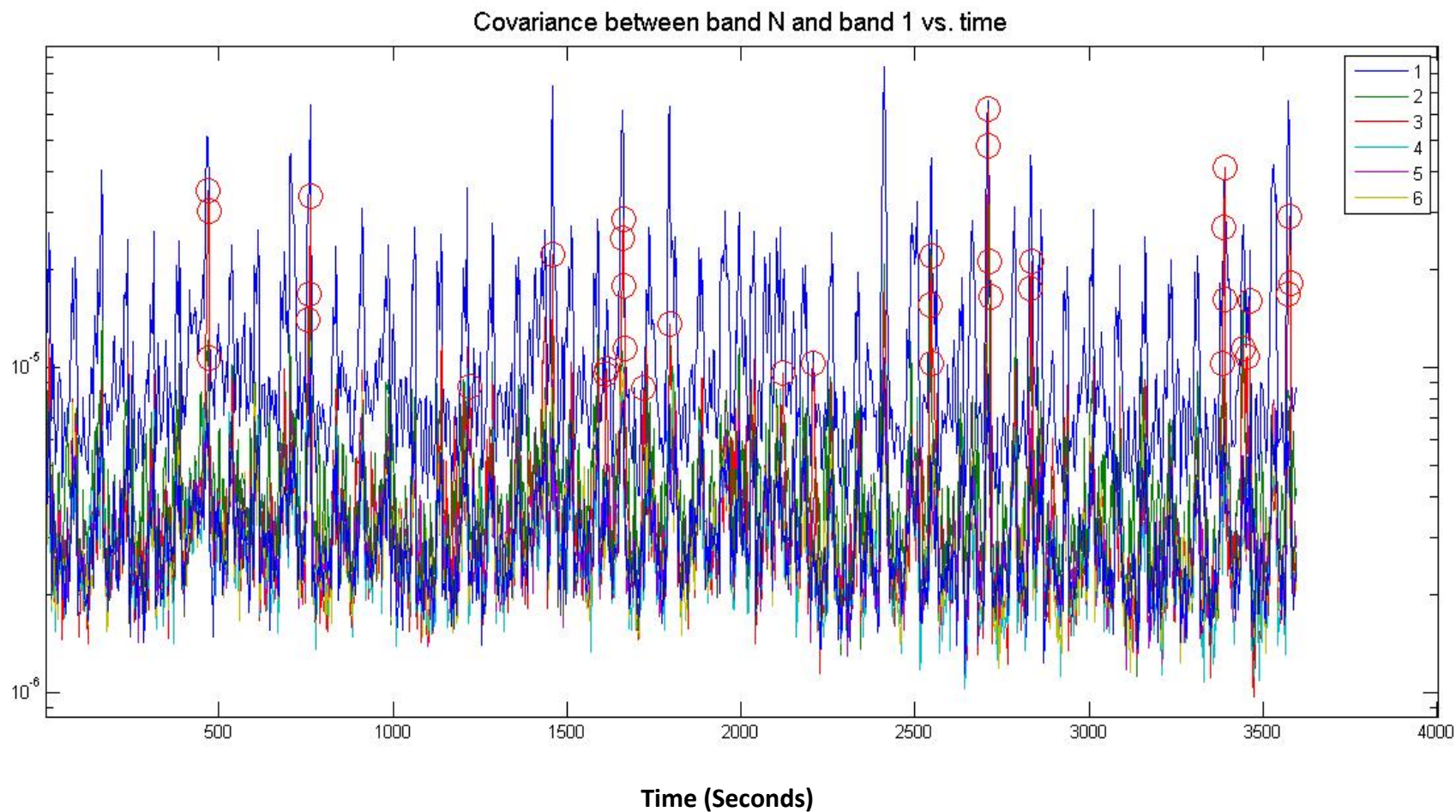
I'd also like to thank Dr. Laela Sayigh of WHOI for sharing the recordings from her collaboration with the Melimoyu Ecosystem Research Institute in Chile and for sharing the manual detections for these recordings, made by Wouter Colpaert.

References

Hoffman, M.D., Garfield, N., and Bland, R.W. (2010), "Frequency synchronization of blue whale calls near Pioneer Seamount", J. Acoust. Soc. Am. 128 (1), 490-494.

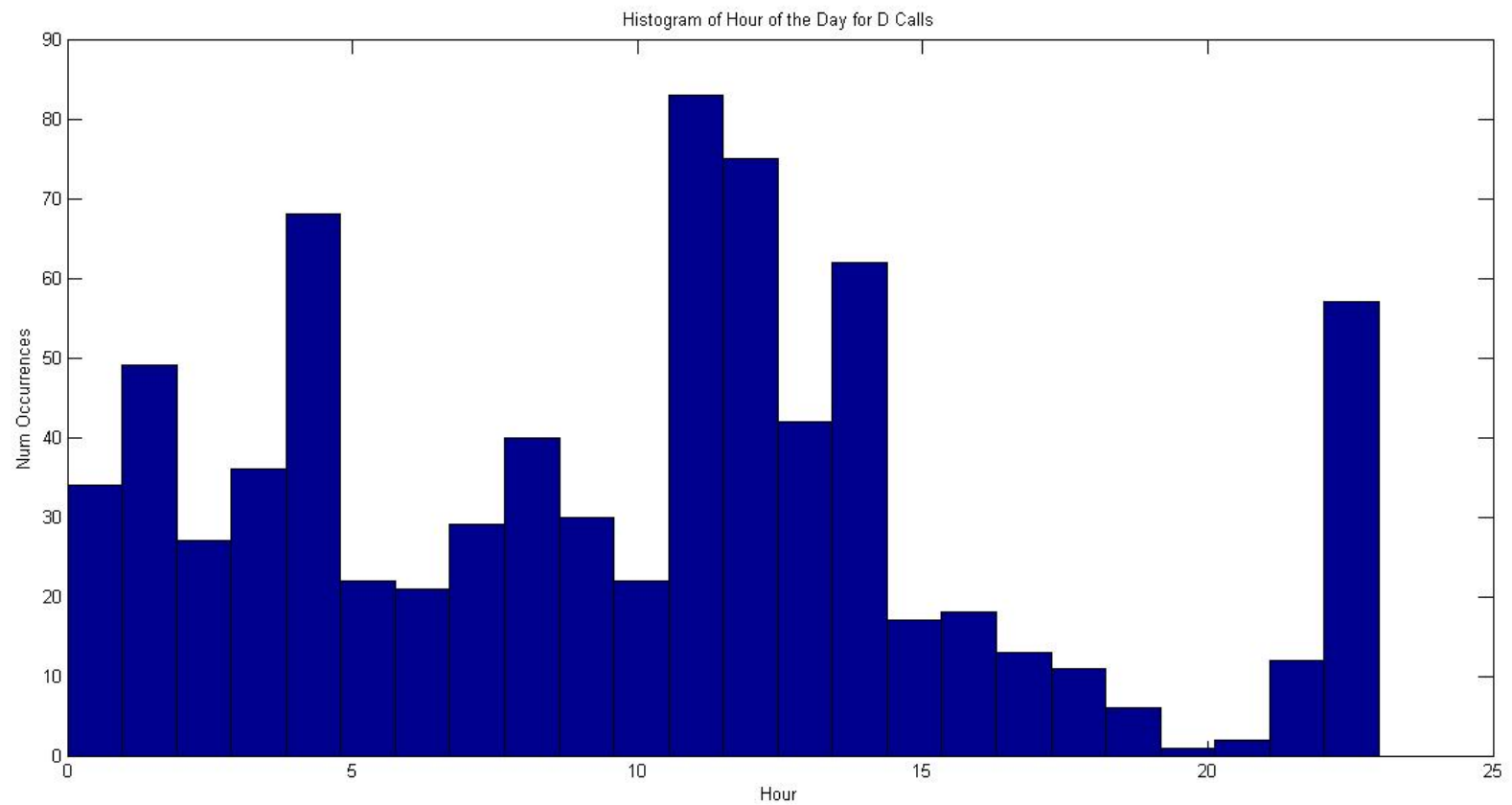
Potamianos, A., and Maragos, P. (1996), "Speech formant frequency and bandwidth tracking using multiband energy demodulation", J. Acoust. Soc. Am. 99 (6), 3795-3806.

Results: Covariance Per Window

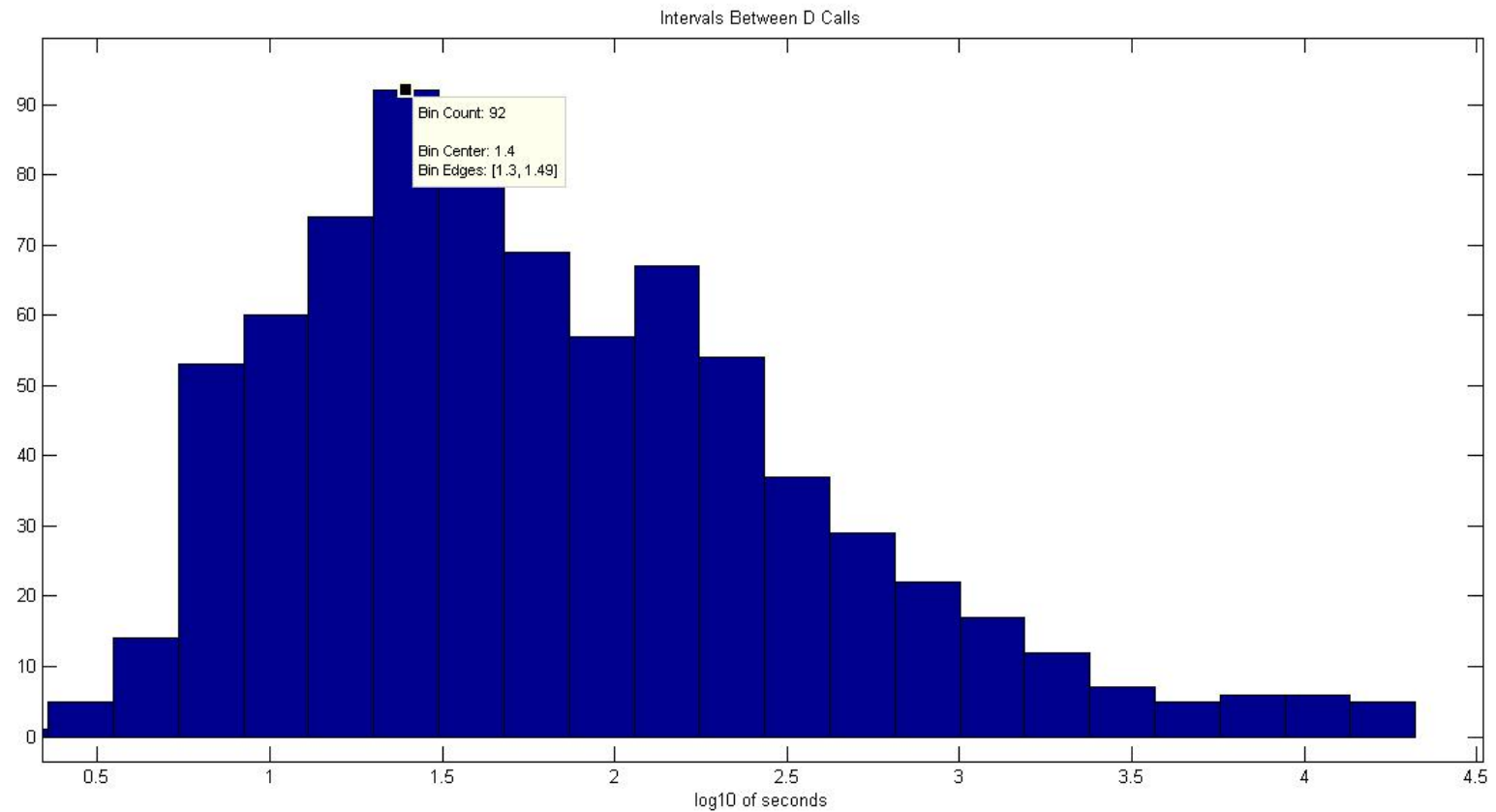


Covariance between demodulated component 3 and demodulated component 1 is in red. Circles indicate 6-second windows in which the signal was detected.

Results: Compare to Distribution of D Calls (Time of Day)



Results: Compare to Distribution of Δt Between D Calls



Peak of distribution is around 25 seconds between D Calls